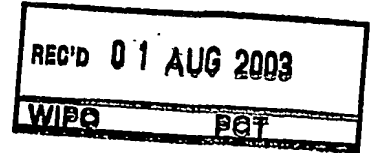


Rec'd PCT/PTO 07 DEC 2004

PCT/JP03/07639

17.06.03

日 本 国 特 許 庁  
JAPAN PATENT OFFICE



別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出 願 年 月 日  
Date of Application: 2002年 6月26日

出 願 番 号  
Application Number: 特願2002-185758  
[ST. 10/C]: [JP2002-185758]

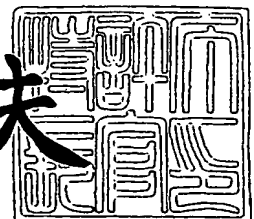
出 願 人  
Applicant(s): 松下電器産業株式会社

PRIORITY DOCUMENT  
SUBMITTED OR TRANSMITTED IN  
COMPLIANCE WITH  
RULE 17.1(a) OR (b)

2003年 7月18日

特許庁長官  
Commissioner,  
Japan Patent Office

今 井 康 夫



出証番号 出証特2003-3057144

【書類名】 特許願

【整理番号】 2022540232

【提出日】 平成14年 6月26日

【あて先】 特許庁長官殿

【国際特許分類】 H04B 1/06

【発明者】

【住所又は居所】 大阪府門真市大字門真1006番地 松下電器産業株式会社内

【氏名】 遠間 正真

【発明者】

【住所又は居所】 大阪府門真市大字門真1006番地 松下電器産業株式会社内

【氏名】 松井 義徳

【発明者】

【住所又は居所】 大阪府門真市大字門真1006番地 松下電器産業株式会社内

【氏名】 能登屋 陽司

【特許出願人】

【識別番号】 000005821

【氏名又は名称】 松下電器産業株式会社

【代理人】

【識別番号】 100097445

【弁理士】

【氏名又は名称】 岩橋 文雄

【選任した代理人】

【識別番号】 100103355

【弁理士】

【氏名又は名称】 坂口 智康

【選任した代理人】

【識別番号】 100109667

【弁理士】

【氏名又は名称】 内藤 浩樹

【手数料の表示】

【予納台帳番号】 011305

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 9809938

【書類名】 明細書

【発明の名称】 多重化装置および逆多重化装置

【特許請求の範囲】

【請求項 1】 音声、動画像、あるいはテキストデータをパケット多重化したデータを出力する多重化装置であって、

前記音声、動画像、あるいはテキストデータを入力する入力手段と、

前記入力手段により入力された前記音声、動画像、あるいはテキストデータを解析して各フレームの表示開始時刻、あるいは前記動画像データのフレームが画面内符号化フレームであるかどうかを示す情報を取得する解析手段と、

前記解析手段により取得した前記各データの表示開始時刻、あるいは前記動画像データが画面内符号化フレームであるかどうかを示す情報に基づいて、前記各データをパケット化するパケット単位決定手段と、

前記パケット単位決定手段により決定されたパケットのヘッダ部を作成するヘッダ部作成手段と、

前記入力手段により入力されたデータから、前記パケット単位決定手段により決定されたパケットのデータ部を作成するデータ部作成手段と、

前記ヘッダ部作成手段により作成されたヘッダ部と、前記データ部作成手段により作成されたデータ部を結合してパケット化する結合手段と、

を備えることを特徴とする多重化装置。

【請求項 2】 前記パケット単位決定手段は、前記パケットに含まれる音声、あるいはテキストの先頭データは、前記パケットに含まれる動画像の先頭データの表示開始時刻以後で、最も近い表示開始時刻をもつデータとなるようにパケット化することを特徴とする請求項 1 記載の多重化装置。

【請求項 3】 前記パケット単位決定手段は、前記パケットに含まれる音声、あるいはテキストの先頭データは、前記パケットに含まれる動画像の先頭データの表示開始時刻以前で、最も近い表示開始時刻をもつデータとなるようにパケット化することを特徴とする請求項 1 記載の多重化装置。

【請求項 4】 前記パケット単位決定手段は、前記動画像データを含むパケットの直前、あるいは直後の前記音声、あるいはテキストデータのパケットの表示

開始時刻は、前記動画像データを含むパケットの表示開始時刻以後で、最も近い表示時刻をもつデータとなるようにパケット化することを特徴とする請求項1記載の多重化装置。

【請求項5】 前記パケット単位決定手段は、前記パケットに含まれる動画像データが1つ以上の画面内符号化フレームのデータを含む際には、前記パケットに含まれる動画像データの先頭に画面内符号化フレームのデータを配置するようにパケット化することを特徴とする請求項1記載の多重化装置。

【請求項6】 音声、動画像、あるいはテキストデータをパケット多重化したデータを入力する逆多重化装置であって、

前記データを入力する入力手段と、

前記入力手段により入力されたデータを解析して、前記パケットのヘッダ部を分離する解析分離手段と、

前記解析分離手段により分離されたヘッダ部を解析して、前記パケットに含まれる音声、動画像、あるいはテキストデータの各フレームの表示開始時刻、サイズ、動画像データのフレームが画面内符号化フレームであるかどうかを示す情報を取得するヘッダ解析手段と、

前記パケット内の前記動画像データに画面内符号化フレームのデータが含まれるかどうかを検索する際には、動画像の先頭データが画面内符号化フレームのデータであるかどうかを前記ヘッダ解析手段の解析結果から判定する検索手段と、

前記検索手段により決定された再生データを取得する取得手段と、

を備えることを特徴とする逆多重化装置。

【請求項7】 音声、動画像、あるいはテキストデータをパケット多重化したデータを出力する多重化方法であって、

前記音声、動画像、あるいはテキストデータを入力する入力ステップと、

前記入力手段により入力された前記音声、動画像、あるいはテキストデータを解析してフレームの表示開始時刻、あるいは前記動画像データのフレームが画面内符号化フレームであるかどうかを示す情報を取得する解析ステップと、

前記解析手段により取得した前記各データの表示開始時刻、あるいは前記動画像データが画面内符号化フレームであるかどうかを示す情報に基づいて、前記各

データをパケット化するパケット単位決定ステップと、

前記パケット単位決定手段により決定されたパケットのヘッダ部を作成するヘッダ部作成ステップと、

前記入力手段により入力されたデータから、前記パケット単位決定手段により決定されたパケットのデータ部を作成するデータ部作成ステップと、

前記ヘッダ部作成手段により作成されたヘッダ部と、前記データ部作成手段により作成されたデータ部を結合してパケット化する結合ステップと、

を備えることを特徴とする多重化方法。

【請求項 8】 前記音声、動画像、あるいはテキストデータを入力する入力ステップと、

前記入力手段により入力された前記音声、動画像、あるいはテキストデータを解析してフレームの表示開始時刻、あるいは前記動画像データのフレームが画面内符号化フレームであるかどうかを示す情報を取得する解析ステップと、

前記解析手段により取得した前記各データの表示開始時刻、あるいは前記動画像データが画面内符号化フレームであるかどうかを示す情報に基づいて、前記各データをパケット化するパケット単位決定ステップと、

前記パケット単位決定手段により決定されたパケットのヘッダ部を作成するヘッダ部作成ステップと、

前記入力手段により入力されたデータから、前記パケット単位決定手段により決定されたパケットのデータ部を作成するデータ部作成ステップと、

前記ヘッダ部作成手段により作成されたヘッダ部と、前記データ部作成手段により作成されたデータ部を結合してパケット化する結合ステップと、

を含むプログラムを記憶した記録媒体。

#### 【発明の詳細な説明】

##### 【0001】

#### 【発明の属する技術分野】

本発明は、音声、動画像、およびテキストデータを多重化する多重化装置、および音声、動画像、およびテキストデータが多重化されたビット列を読み込んで逆多重化する逆多重化装置に関するものである。

## 【0002】

## 【従来の技術】

近年、通信ネットワークの大容量化および伝送技術の進歩により、インターネット上でのPC向け動画配信サービスが普及してきた。さらに、無線端末上での動画配信についても、無線網における受信端末の規格を定める国際標準化団体である3GPP (Third Generation Partnership Project) における規格としてTS26.234 (Transparent end-to-end packet switched streaming service) が定められるなど、携帯端末においても動画配信サービスの拡大が見込まれる。音声、動画、静止画、およびテキストなどのメディアデータを蓄積、および配信する際には、メディアデータの再生に必要なヘッダ情報とメディアデータを多重化するが、そのための多重化ファイルフォーマットとしてMP4が標準化された。MP4は、ISO/IEC JTC1/SC29/WG11 (International Standardisation Organization/International Engineering Consortium) において標準化された多重化ファイルフォーマットであり、3GPPのTS26.234においても採用されていることから、今後普及していくと考えられる。

## 【0003】

ここで、MP4のデータ構造について説明する。MP4では、ヘッダ情報やメディアデータはBoxと呼ばれるオブジェクト単位で格納される。図8は、Boxの構造を説明するための図である。Boxは、以下のフィールドを持つ。

## 【0004】

size: sizeフィールドも含めたBox全体のサイズ

type: Boxの識別子であり、通常はアルファベット4文字で表される

version: Boxのバージョン番号

flags: Box毎に設定されるフラグ情報

データ: ヘッダ情報やメディアデータが格納される。

## 【0005】

ただし、`version`と`flags`は必須でないため、`Box`によってはこれらのフィールドは存在しない。以後、`Box`の参照には`type`フィールドの識別子を使用することとし、例えば、`type`が`'moov'`である`Box`は、`moov`と呼ぶ。MP4ファイルでは、使用が必須である基本部に加えて、拡張部を使用することができる。なお、拡張部については現在標準化中である。

#### 【0006】

まず、基本部の構造について説明する。図9は、MP4ファイルの基本部を説明するための図である。MP4ファイルは、`ftyp`、`moov`、`mdat`の3つの基本`Box`から構成され、`ftyp`がファイルの先頭に配置される。

#### 【0007】

`ftyp`は、MP4ファイルを識別するための情報を含む。MP4に格納するメディアデータおよびメディアデータの符号化形式については、標準化団体やサービス事業者が独自に規定できたため、MP4ファイルがどの規定に従ったものであるかを、`ftyp`により識別する。

#### 【0008】

`mdat`には、サンプルと呼ばれる単位でメディアデータを格納する。サンプルとは、MP4でメディアデータを扱う際の最小単位であり、オーディオのフレームやMPEG-4 VisualのVOP (Video Object Plane) に相当する。また、`mdat`に含まれる各メディアのデータはトラックと呼ばれる。

#### 【0009】

`moov`には、`mdat`に含まれるサンプルのヘッダ情報が格納される。図10は、`moov`の構造を説明するための図である。`moov`内では`Box`が階層的に配置されており、`trak`と呼ばれる`Box`を用いてオーディオ、ビデオなどトラック毎のヘッダ情報が格納される。さらに、下位の`Box`においてトラックのサンプル単位のヘッダ情報を格納する。

#### 【0010】

基本部ではトラック全体に対してヘッダ情報が付加されるが、拡張部においては、トラックを分割し、分割したトラックのそれぞれに対してヘッダ情報が付加



される。ただし、拡張部を使用する際にも `f t y p` と `m o o v` は必須であり、符号化形式などトラック全体に共通なヘッダ情報は `m o o v` に格納される。

#### 【0011】

図11に拡張部を含むMP4ファイルの構造を示す。分割されたトラックのヘッダ情報はそれぞれ `m o o f` に格納され、`m o o f` に対して `m d a t` が付加される。図12は、`m o o f` の構造を説明するための図である。`m o o f` においても、`m o o v` と同様に `B o x` が階層的に配置される。`m o o f` 直下には、`m o o f` 全体のヘッダ情報を格納する `m f h d` と、トラック毎のヘッダ情報を格納する `t r a f` が配置される。`t r a f` は、`m o o v` における `t r a k` と同様に、単一トラックのヘッダ情報を格納するが、複数の `t r a f` を使用してヘッダ情報を格納してもよい。ただし、ヘッダ情報を複数の `t r a f` に分割して格納する際には、`t r a f` に含まれるサンプルの再生時間が早い順に、`t r a f` を配置する。`t r a f` には、`t f h d` と1つ以上の `t r u n` が格納され、これらの `B o x` にサンプル単位のヘッダ情報が含まれる。`t f h d` は、`t r a f` が格納するトラックの識別子（トラックID）、およびヘッダ情報のデフォルト値を示し、`t r u n` は、サンプル単位のヘッダ情報を記述する。図18に `t r u n` の構造を示す。各フィールドの内容を以下で説明する。

#### 【0012】

- ・ `t r _ f l a g s` : `d a t a _ o f f s e t`、`f i r s t _ s a m p l e _ f l a g s`、`s a m p l e _ d u r a t i o n`、`s a m p l e _ s i z e`、`s a m p l e _ f l a g s`、`s a m p l e _ c o m p o s i t i o n _ t i m e _ o f f s e t` の各フィールドが存在するかどうかを示すフラグ
- ・ `s a m p l e _ c o u n t` : `t r u n` にヘッダ情報が格納されるサンプル数
- ・ `d a t a _ o f f s e t` : `t r u n` に含まれる先頭サンプルのデータが `m d a t` のどこに格納されるかを示すポインタ
- ・ `f i r s t _ s a m p l e _ f l a g s` : `t r u n` に含まれる先頭サンプルのみ、`s a m p l e _ f l a g s` の値が異なる際に、先頭サンプルの `s a m p l e _ f l a g s` の値を `f i r s t _ s a m p l e _ f l a g s` で上書きすることができる。

## 【0013】

次に、以下の4つのフィールドがテーブルのエントリーとなり、各エントリーがサンプル毎の情報を示す。なお、エントリーはsample\_\_countで示される個数だけ存在し、いずれのフィールドが存在するかはtr\_\_flagsにより示される。また、フィールドが存在しないときには、tfhdあるいは、moov内のmvexと呼ばれるボックスに示される各フィールドのデフォルト値を使用するものとする。

## 【0014】

- ・sample\_\_duration: サンプルの表示時間長
- ・sample\_\_size: サンプルのサイズ
- ・sample\_\_flags: サンプルがランダムアクセス可能であるかどうかを示すフラグ
- ・sample\_\_composition\_\_time: サンプルのデコード時刻と表示時刻との差分値

なお、trunには再生時間の早いサンプルから順にヘッダ情報が記述される。このため、サンプルのヘッダ情報を検索する際には、先頭のtrafから順にtfhd内のトラックIDを参照することにより、取得したいトラックのヘッダ情報を含むtrafを検索していく。なお、traf内でサンプルのヘッダ情報を検索する際には、先頭のtrunから順にサンプルのヘッダ情報を検索する。

## 【0015】

図13は、拡張部の使用例を説明するための図である。ここでは、コンテンツの格納方法について2通りの例を示す。なお、コンテンツの再生時間は60秒であるとする。

## 【0016】

MP4ファイル1は、基本部と拡張部の両方を使用してメディアデータを格納している。コンテンツは、(1) 0-30秒、(2) 30-45秒、(3) 45-60秒の3つの部分に分割される。(1)の部分は基本部を用いてmdat1に格納され、(2)および(3)の部分は拡張部を用いてそれぞれmdat2とmdat3に格納される。拡張部のBoxの配置は、moof1、mdat2、

m o o f 2、m d a t 3の順になっており、m d a t 2のヘッダ情報はm o o f 1に、m d a t 3のヘッダ情報はm o o f 2により示される。

#### 【0017】

MP4ファイル2は、拡張部のみを使用してメディアデータを格納している。コンテンツは(1)0-30秒、(2)30-60秒の2つの部分に分割され、それぞれm d a t 4とm d a t 5に格納される。拡張部のB o xの配置は、m o o f 3、m d a t 4、m o o f 4、m d a t 5の順になっており、m d a t 4のヘッダ情報はm o o f 3に、m d a t 5のヘッダ情報はm o o f 4により示される。基本部には、f t y pとm o o vのみが含まれ、m d a tは含まれない。

#### 【0018】

以下、MP4ファイルの拡張部を作成する方法の従来例について説明する。図14は、従来の多重化装置の構成を示すブロック図である。ここでは、ビデオとオーディオのデータを多重化するものとする。ビデオの入力データs t r e a m 10は、入力手段I n p u t 1に入力される。入力手段I n p u t 1は、入力データs t r e a m 10を、入力データs t r e a m 11としてメモリM e m 1に入力する。メモリM e m 1は、入力データs t r e a m 11を格納する。解析手段P a r s e 1は、メモリM e m 1から1サンプル分のビデオデータを順に取得、解析し、サンプルのヘッダ情報s p l I n f o 1をパケット単位決定手段P a c k t i z e 1に入力する。入力手段I n p u t 2は、入力データs t r e a m 20を、入力データs t r e a m 21としてメモリM e m 2に入力する。メモリM e m 2は、入力データs t r e a m 21を格納する。解析手段P a r s e 2は、メモリM e m 2から1サンプル分のオーディオデータを順に取得、解析し、サンプルのヘッダ情報s p l I n f o 2をパケット単位決定手段P a c k t i z e 1に入力する。

#### 【0019】

なお、ヘッダ情報s p l I n f o 1、およびs p l I n f o 2には、サンプルのサイズ、再生時間長、および、ビデオの場合にはサンプルが画面内符号化フレームかどうかを示す情報が含まれる。パケット単位決定手段P a c k t i z e 1は、パケットに含まれるサンプル数が一定となるように、ビデオとオーディオの

パケット単位を決定するとともに、各パケットのヘッダ情報を作成する。

#### 【0020】

図15は、パケット単位決定手段Packetize1の動作を示すフローチャートである。なお、Nは1パケットに格納するサンプルの数を示し、値は予め定められているものとする。なお、パケットの先頭サンプルの番号をfirst\_\_splとする。step101において、i番目のサンプルのヘッダ情報splInfoが入力される。step102において、サンプルのヘッダ情報splInfoを取得し、テーブル情報pktInfoをメモリMem3に出力する。テーブル情報pktInfoにはサンプルのサイズ、再生時間長、および、ビデオの場合にはサンプルが画面内符号化フレームかどうかを示す情報が含まれ、Mem3ではこれらのサンプル単位の情報がテーブルに格納される。

#### 【0021】

図16に、ビデオサンプルのヘッダ情報を格納するテーブルの一例を示す。サンプル毎にサイズ、再生時間長、画面内符号化フレームであるかどうかを示す情報が格納され、パケット内の先頭サンプルは、サイズが300バイト、再生時間長が30ms、かつ画面内符号化フレームであることが示される。2番目以降のサンプルの情報は、順にテーブルに追加される。

#### 【0022】

step103において、パケットに含まれる総サンプル数num\_\_splに1を加算する。なお、num\_\_splの初期値は0であるものとする。次に、step104においてnum\_\_splとNを比較し、num\_\_splがNよりも小さい場合には、step105においてiに1を加算し、step101の処理に戻る。num\_\_splがN以上である際には、step106においてfirst\_\_splからi番目サンプルまでをパケット化するように決定する。step101からstep106までの処理を、ビデオとオーディオについて行う。オーディオとビデオのパケット単位を決定したのち、パケット単位決定手段Packetize1は、パケット作成信号pktsignal1をヘッダ部作成手段moofMakeに入力する。

#### 【0023】

ヘッダ部作成手段Moo f Makeは、パケット作成信号p k t s i g n a l 1の入力後、パケットに含まれるサンプルのヘッダ情報p k t D a tをメモリM e m 3から取得し、m o o fを作成する。さらに、m o o f作成後、m o o fデータm o o f D a tを結合手段M u xに入力し、m d a tの作成に必要なMDA T情報m d a t I n f oをデータ部作成手段M d a t M a k eに入力する。

#### 【0024】

なお、MDA T情報m d a t I n f oには、パケット内のサンプルのデータがメモリM e m 1、およびメモリM e m 2のどこに格納されているかを示すポインタ情報が含まれる。データ部作成手段M d a t M a k eは、メモリM e m 1、およびメモリ2からパケットに含まれるサンプルのオーディオ、およびビデオデータを取得してm d a tの作成を行い、m d a tデータm d a t D a tを結合手段M u xに入力する。結合手段M u xは、m o o fデータm o o f D a tとm d a tデータm d a t D a tとを結合してMP 4の拡張部データm p 4を出力する。

#### 【0025】

##### 【発明が解決しようとする課題】

このような従来の多重化装置により多重化されたMP 4ファイルの拡張部を再生する際には、以下のような課題がある。

#### 【0026】

従来の多重化装置では、パケットに含まれるサンプルの表示開始時刻を考慮せずに多重化を行っていたため、ある表示開始時刻に対応するサンプルの格納位置がメディア毎に大きく異なることがあり、再生時のデータアクセスの効率が悪いという第1の課題があった。前記第1の課題について、図17の(1)を用いて説明する。なお、MP 4ファイル内にビデオとオーディオが多重化されており、各m o o fに含まれるサンプルのデータは、直後のm d a tに格納されているものとする。

#### 【0027】

表示開始時刻が20sであるビデオサンプルは、m o o f 1の先頭サンプルとして格納され、表示開始時刻が20sであるオーディオサンプルは、m o o f 10の最終サンプルとして格納される。このため、コンテンツにおいて再生時間が

20sである部分を再生するには、ビデオサンプルのヘッダ情報を取得してから、オーディオサンプルのヘッダ情報を取得するまでにmoo f 1からmoo f 10までを検索する必要がある、サンプルの検索に多くの計算を要する。

#### 【0028】

また、従来の多重化装置では、ビデオの画面内符号化フレームに相当するサンプルをパケット内の何番目のサンプルとするか考慮せずに多重化を行っていたため、ランダムアクセス可能なサンプルを検索する際には、パケットの最終サンプルまで検索する必要がある、検索に必要な計算量が多いという第2の課題があった。前記第2の課題について、図17の(2)を用いて説明する。この例では、ビデオの画面内符号化フレームに対応するサンプルの内、i番目、およびi+1番目のサンプルがともにパケットの最終サンプルとして格納されている。このため、ランダムアクセス時にパケットの最終サンプルまで検索する必要がある、効率が悪い。

#### 【0029】

##### 【課題を解決するための手段】

前記課題を解決するため、本発明の多重化装置においては、moo fに含まれる各メディアのサンプルの表示開始時刻が一定となるように多重化を行う、あるいは、moo fに画面内符号化フレームのサンプルを格納する際には、moo f内の先頭サンプルとして格納するようにmoo fを作成するものであり、以下のような手段で構成される。

#### 【0030】

本発明の請求項1にかかる多重化装置は、音声、動画像、あるいはテキストデータをパケット多重化したデータを出力する多重化装置であって、前記音声、動画像、あるいはテキストデータを入力する入力手段と、前記入力手段により入力された前記音声、動画像、あるいはテキストデータを解析して各フレームの表示開始時刻、あるいは前記動画像データのフレームが画面内符号化フレームであるかどうかを示す情報を取得する解析手段と、前記解析手段により取得した前記各データの表示開始時刻、あるいは前記動画像データが画面内符号化フレームであるかどうかを示す情報に基づいて、前記各データをパケット化するパケット単位

決定手段と、前記パケット単位決定手段により決定されたパケットのヘッダ部を作成するヘッダ部作成手段と、前記入力手段により入力されたデータから、前記パケット単位決定手段により決定されたパケットのデータ部を作成するデータ部作成手段と、前記ヘッダ部作成手段により作成されたヘッダ部と、前記データ部作成手段により作成されたデータ部を結合してパケット化する結合手段と、を備えるものである。

#### 【0031】

本発明の請求項2にかかる多重化装置は、請求項1の多重化装置において、前記パケット単位決定手段は、前記パケットに含まれる音声、あるいはテキストの先頭データは、前記パケットに含まれる動画像の先頭データの表示開始時刻以後で、最も近い表示開始時刻をもつデータとなるようにパケット化することを特徴とするものである。

#### 【0032】

本発明の請求項3にかかる多重化装置は、請求項1の多重化装置において、前記パケット単位決定手段は、前記パケットに含まれる音声、あるいはテキストの先頭データは、前記パケットに含まれる動画像の先頭データの表示開始時刻以前で、最も近い表示開始時刻をもつデータとなるようにパケット化することを特徴とするものである。

#### 【0033】

本発明の請求項4にかかる多重化装置は、請求項1の多重化装置において、前記パケット単位決定手段は、前記動画像データを含むパケットの直前、あるいは直後の前記音声、あるいはテキストデータのパケットの表示開始時刻は、前記動画像データを含むパケットの表示開始時刻以後で、最も近い表示時刻をもつデータとなるようにパケット化することを特徴とするものである。

#### 【0034】

本発明の請求項5にかかる多重化装置は、請求項1の多重化装置において、前記パケット単位決定手段は、前記パケットに含まれる動画像データが1つ以上の画面内符号化フレームのデータを含む際には、前記パケットの先頭に画面内符号化フレームのデータを配置するようにパケット化することを特徴とするものである。

る。

#### 【0035】

本発明の請求項6にかかる逆多重化装置は、音声、動画像、あるいはテキストデータをパケット多重化したデータを入力する逆多重化装置であって、前記データを入力する入力手段と、前記入力手段により入力されたデータを解析して、前記パケットのヘッダ部を分離する解析分離手段と、前記解析分離手段により分離されたヘッダ部を解析して、前記パケットに含まれる音声、動画像、あるいはテキストデータの各フレームの表示開始時刻、サイズ、動画像データのフレームが画面内符号化フレームであるかどうかを示す情報を取得するヘッダ解析手段と、前記パケット内の前記動画像データに画面内符号化フレームのデータが含まれるかどうか検索する際には、動画像の先頭データが画面内符号化フレームのデータであるかどうかを前記ヘッダ解析手段の解析結果から判定する検索手段と、前記検索手段により決定された再生データを取得する取得手段と、を備えるものである。

#### 【0036】

本発明の請求項7にかかる多重化方法は、音声、動画像、あるいはテキストデータをパケット多重化したデータを出力する多重化方法であって、前記音声、動画像、あるいはテキストデータを入力する入力ステップと、前記入力手段により入力された前記音声、動画像、あるいはテキストデータを解析してフレームの表示開始時刻、あるいは前記動画像データのフレームが画面内符号化フレームであるかどうかを示す情報を取得する解析ステップと、前記解析手段により取得した前記各データの表示開始時刻、あるいは前記動画像データが画面内符号化フレームであるかどうかを示す情報に基づいて、前記各データをパケット化するパケット単位決定ステップと、前記パケット単位決定手段により決定されたパケットのヘッダ部を作成するヘッダ部作成ステップと、前記入力手段により入力されたデータから、前記パケット単位決定手段により決定されたパケットのデータ部を作成するデータ部作成ステップと、前記ヘッダ部作成手段により作成されたヘッダ部と、前記データ部作成手段により作成されたデータ部を結合してパケット化する結合ステップと、を備えるものである。



## 【0037】

## 【発明の実施の形態】

## (実施の形態1)

本発明の多重化装置の実施の形態1について、図1から図3を参照しながら説明する。なお、ここではビデオとオーディオのデータを多重化するものとする。

## 【0038】

ビデオの入力データstream10は、入力手段Input1に入力される。入力手段Input1は、入力データstream10を、入力データstream11としてメモリMem1に入力する。メモリMem1は、入力データstream11を格納する。解析手段Parse1は、メモリMem1から1サンプル分のビデオデータを順に取得、解析し、サンプルのヘッダ情報splInfo1をパケット単位決定手段Packetize2に入力する。

## 【0039】

入力手段Input2は、入力データstream20を、入力データstream21としてメモリMem2に入力する。メモリMem2は、入力データstream21を格納する。解析手段Parse2は、メモリMem2から1サンプル分のオーディオデータを順に取得、解析し、サンプルのヘッダ情報splInfo2をパケット単位決定手段Packetize2に入力する。なお、ヘッダ情報splInfo1、およびヘッダ情報splInfo2には、サンプルのサイズ、再生時間長、および、ビデオの場合にはサンプルが画面内符号化フレームに相当するかどうかを示す情報が含まれる。パケット単位決定手段Packetize2は、まずビデオのパケット化単位を決定し、次にオーディオのパケット化単位を決定する。

## 【0040】

図2は、パケット単位決定手段Packetize2においてビデオのパケット化単位を決定する際の動作を示すフローチャートである。なお、パケットに含まれるサンプルの総再生時間の目標値をdeltaとし、予め決められた値を用いるものとする。例えば、1パケットに5s分のサンプルを格納したいときには、deltaを5sに設定する。なお、パケットの先頭サンプルの番号をfir

s t \_ s p l とし、パケットの再生終了時間を示すパラメータである e n d \_ t i m e の初期値は、d e l t a であるものとする。

【0041】

s t e p 2 0 1 において、i 番目のサンプルのヘッダ情報 s p l I n f o が入力される。s t e p 2 0 2 において、サンプルのヘッダ情報 s p l I n f o を取得し、テーブル情報 p k t I n f o をメモリ M e m 3 に出力する。テーブル情報 p k t I n f o にはサンプルのサイズ、再生時間長、および、ビデオの場合にはサンプルが画面内符号化フレームかどうかを示す情報が含まれ、M e m 3 ではこれらのサンプル単位の情報がテーブルに格納される。

【0042】

s t e p 2 0 3 において、i 番目サンプルの再生終了時間 s p l \_ t i m e と e n d \_ t i m e とを比較し、s p l \_ t i m e が e n d \_ t i m e よりも小さい場合には、s t e p 2 0 4 において i に 1 を加算し、s t e p 2 0 1 の処理に戻る。s p l \_ t i m e が e n d \_ t i m e 以上である際には、s t e p 2 0 5 において s p l \_ t i m e に d e l t a を加算することにより次パケットの再生終了時間を決定する。なお、次のパケットの終了時間は、e n d \_ t i m e に d e l t a を加算することにより求めてもよい。

【0043】

次に、s t e p 2 0 6 において f i r s t \_ s p l から i 番目サンプルまでをパケット化するように決定する。なお、s p l \_ t i m e が e n d \_ t i m e よりも大きい場合には、f i r s t \_ s p l から i - 1 番目のサンプルまでをパケット化することにより、パケットの再生時間が、再生時間の目標値以下で最大となるようにしてもよい。このとき、次パケットの再生終了時間には、i - 1 番目のサンプルの再生終了時間に d e l t a を加算した値を用いる。

【0044】

次に、オーディオのパケット化単位を決定する。オーディオの k 番目パケット化単位を決定する方法を以下に示す。ビデオの k 番目パケット化単位の表示開始時刻より後で、最も早い表示開始時刻をもつサンプルを k 番目パケット化単位の先頭サンプルとし、ビデオの k + 1 番目パケット化単位の表示開始時刻より後で

、最も早い表示開始時刻をもつサンプルをk番目パケット化単位の最終サンプルとする。なお、ビデオのk番目パケット化単位の表示開始時刻より前で、最も遅い表示開始時刻をもつサンプルをk番目パケット化単位の先頭サンプルとし、ビデオのk+1番目パケット化単位の表示開始時刻より前で、最も遅い表示開始時刻をもつサンプルをk番目パケット化単位の最終サンプルとしてもよい。オーディオのパケット化単位決定後、パケット作成信号p k t s i g n a lをヘッダ部作成手段m o o f M a k eに入力する。

#### 【0045】

ヘッダ部作成手段M o o f M a k eは、パケット作成信号p k t s i g n a lの入力後、パケットに含まれるサンプルのヘッダ情報p k t D a tをメモリM e m 3から取得し、m o o fを作成する。なお、AMRのように、データの途中でモードの切り替わりが発生する符号化方式により符号化されたメディアデータを扱う際には、モードの切り替わりに応じてt r a fを切り替えてもよい。さらに、m o o f作成後、m o o fデータm o o f D a tを結合手段M u xに入力し、m d a tの作成に必要なMDAT情報m d a t I n f oをデータ部作成手段M d a t M a k eに入力する。なお、MDAT情報m d a t I n f oには、パケット内のサンプルのデータがメモリM e m 1、およびメモリM e m 2のどこに格納されているかを示すポインタ情報が含まれる。

#### 【0046】

データ部作成手段M d a t M a k eは、メモリM e m 1、およびメモリM e m 2からパケットに含まれるサンプルのオーディオ、およびビデオデータを取得してm d a tの作成を行い、m d a tデータm d a t D a tを結合手段M u xに入力する。結合手段M u xは、m o o fデータm o o f D a tとm d a tデータm d a t D a tとを結合してMP4の拡張部データm p 4を出力する。

#### 【0047】

図3は、拡張部データm p 4のデータ構造を示す。図3の(1)では、パケットは、m o o fとm d a tから構成され、各m o o fにビデオとオーディオのサンプルのヘッダ情報が格納され、サンプルのデータがm d a tに格納される。パケット1に含まれるビデオとオーディオの先頭サンプルの表示開始時刻はともに

20 s であり、パケット 2 に含まれるビデオとオーディオの先頭サンプルの表示開始時刻はともに 30 s である。

#### 【0048】

このように格納することにより、同一パケット内でのオーディオとビデオの表示開始時刻が揃うため、再生時にサンプルを検索する際の計算量を削減することができる。また、任意のパケット単位でデータを分割しても各メディアの再生終了時間が揃っているため、パケット単位でデータを分割することにより、データサイズを制限することができるという利点もある。

#### 【0049】

なお、図 3 の (2) のように格納することも可能である。この例では、パケットは、m o o f と m d a t から構成され、パケットにはビデオ、またはオーディオのどちらか一方のデータが格納される。パケット 1 はビデオデータを含み、先頭サンプルの表示開始時刻は 20 s であり、パケット 2 はオーディオデータを含み、先頭サンプルの表示開始時刻は 20 s である。以降のパケットにおいても、オーディオパケットの表示開始時刻が、直前のビデオパケットの表示開始時刻付近となるようにパケット化される。

#### 【0050】

##### (実施の形態 2)

本発明の多重化装置の実施の形態 2 について、図 4 から図 6 を参照しながら説明する。

#### 【0051】

多重化装置の構成は実施の形態 1 に係る多重化装置と同一であり、パケット単位決定手段 P a c k t i z e 2 の動作が異なることから、ここではパケット単位決定手段 P a c k t i z e 2 の動作について説明する。なお、ビデオとオーディオのデータを多重化するものとする。

#### 【0052】

パケット単位決定手段 P a c k t i z e 2 では、まずビデオのパケット化単位を決定し、次にオーディオのパケット化単位を決定する。図 4 は、ビデオデータが入力された際のパケット単位決定手段 P a c k t i z e 2 の動作を示すフロー

チャートである。なお、パケットに含まれるサンプルの総再生時間の目標値をTとし、予め決定された値を用いるものとする。

#### 【0053】

例えば、1パケットに5 s分のサンプルを格納したいときには、Tを5 sに設定する。なお、packet\_durは、パケットに含まれるサンプルの総再生時間を示し、初期値は0であるものとする。なお、パケットの先頭サンプルの番号をfirst\_splとする。step 301において、i番目のサンプルのヘッダ情報spl Infoが入力される。step 302において、サンプルのヘッダ情報spl Infoを取得し、テーブル情報pkt InfoをメモリMem2に出力する。ヘッダ情報pkt Infoにはサンプルのサイズ、再生時間長、および、サンプルが画面内符号化フレームに相当するかどうかを示す情報が含まれ、Mem2ではこれらのサンプル単位の情報がテーブルに格納される。

#### 【0054】

step 303において、i番目のサンプルが画面内符号化フレームであるかどうか判定する。画面内符号化フレームでない場合は、step 305においてpacket\_durにi番目サンプルの再生時間長を加算し、step 306においてiに1を加算した後に、step 301の処理に戻る。画面内符号化フレームである際は、step 304においてpacket\_durとTを比較し、packet\_durがTよりも小さい場合には、step 305の処理に戻る。packet\_durがT以上である際には、step 307においてfirst\_splからi-1番目サンプルまでをパケット化するように決定し、パケット作成信号pkt signal 2をヘッダ部作成手段mo of Makeに入力する。

#### 【0055】

なお、次パケットの先頭サンプルは、i番目のサンプルとなる。こうすることにより、パケットに画面内符号化フレームが含まれるときには、パケットの先頭サンプルは画面内符号化フレームのデータとなり、パケット単位でランダムアクセスすることができる。さらに、先頭trunの先頭サンプルがランダムアクセス可能であることを示す際には、first\_sample\_flagsフィー

ルドを使用することにより `sample_flags` フィールドを省略することができるので、オーバーヘッドの削減も可能となる。ただし、本方法では、画面内符号化フレームの間隔の増加すると、パケットに含まれるサンプル数、あるいはパケットの再生時間長が増加する。そこで、第2の方法として、パケットに含まれるサンプルの総再生時間が目標値 `T` 以下となるようにパケット化を行ってもよい。

#### 【0056】

図5は、第2の方法における、ビデオデータが入力された際のパケット単位決定手段 `Packtiz2` の動作を示すフローチャートである。`step311` において、`i` 番目のサンプルのヘッダ情報 `splInfo` が入力される。`step312` において、サンプルのヘッダ情報 `splInfo` を取得し、テーブル情報 `pktInfo` をメモリ `Mem2` に出力する。ヘッダ情報 `pktInfo` にはサンプルのサイズ、再生時間長、および、サンプルが画面内符号化フレームに相当するかどうかを示す情報が含まれ、`Mem2` ではこれらのサンプル単位の情報がテーブルに格納される。

#### 【0057】

`step313` において、`packet_dur` と `T` を比較する。`packet_dur` が `T` よりも小さい場合には、`step314` においてサンプルが画面内符号化フレームであるかどうか判定する。画面内符号化フレームでない場合は、`step315` において `packet_dur` に `i` 番目サンプルの再生時間長を加算し、`step316` において `i` に1を加算した後に、`step311` の処理に戻る。画面内符号化フレームである際には、`step317` において `first_spl` から `i-1` 番目サンプルまでをパケット化するように決定し、パケット作成信号 `pktsignal2` をヘッダ部作成手段 `moofMake` に入力する。

#### 【0058】

なお、`step314` において、`packet_dur` が予め定めた値以上である場合にのみ `step317` に進み、そうでないときは `step315` に戻ることにしてもよい。`step313` において、`packet_dur` が `T` 以上で

あると判定された際には、step 317 の処理を行う。なお、次パケットの先頭サンプルは、i 番目のサンプルとなる。

#### 【0059】

次に、オーディオのパケット化単位を決定する。オーディオの k 番目パケット化単位を決定する方法を以下に示す。ビデオの k 番目パケット化単位の表示開始時刻より後で、最も早い表示開始時刻をもつサンプルを k 番目パケット化単位の先頭サンプルとし、ビデオの k + 1 番目パケット化単位の表示開始時刻より後で、最も早い表示開始時刻をもつサンプルを k 番目パケット化単位の最終サンプルとする。なお、ビデオの k 番目パケット化単位の表示開始時刻より前で、最も遅い表示開始時刻をもつサンプルを k 番目パケット化単位の先頭サンプルとし、ビデオの k + 1 番目パケット化単位の表示開始時刻より前で、最も遅い表示開始時刻をもつサンプルを k 番目パケット化単位の最終サンプルとしてもよい。オーディオのパケット化単位決定後、パケット作成信号 `pkt signal` をヘッダ部作成手段 `moof Make` に入力する。

#### 【0060】

図 6 は、拡張部データ `mp4` のデータ構造を示す。図 6 の (1) では、パケットは、`moof` と `mdat` から構成され、各 `moof` にビデオとオーディオのサンプルのヘッダ情報が格納され、サンプルのデータが `mdat` に格納される。パケット 1、およびパケット 2 に含まれるビデオの先頭サンプルは画面内符号化フレームである。オーディオは、任意のサンプルに対してランダムアクセスできるため、パケットの先頭サンプルから再生を開始することができ、ランダムアクセス可能なサンプルの検索に要する計算量を削減することができる。

#### 【0061】

なお、図 6 (2) のように格納することも可能である。この例では、パケットは、`moof` と `mdat` から構成され、パケットにはビデオ、またはオーディオのどちらか一方のデータが格納される。パケット 1 はビデオデータを含み、先頭サンプルは画面内符号化フレームである。パケット 2 はオーディオデータを含み、先頭サンプルの表示開始時刻はパケット 1 の表示開始時刻付近となる。以降のパケットについても同様にパケット化される。

## 【0062】

## (実施の形態3)

本発明の逆多重化装置の実施の形態3について、図7を参照しながら説明する。図7は、逆多重化装置の構成を示すブロック図である。MP4ファイルデータmp4は、入力手段Inputに入力される。入力手段Inputは、MP4ファイルデータmp4を、MP4データmp4Dat1としてメモリMem1に入力する。メモリMem1は、MP4データmp4Dat1を格納する。MP4解析分離手段rootDemuxは、メモリMem1からMP4データmp4Dat2を取得して解析し、moovデータMoovとmoofデータMoofを分離する。分離されたmoovデータMoovはmoov解析手段moovDemuxに、moofデータMoofはmoof解析手段moofDemuxに入力される。moov解析手段moovDemuxは、moovデータMoovを解析してタイムスケールや符号化パラメータなどメディアデータの解析に必要なメディア情報MoovInfを取得し、moof解析手段moofDemuxに入力する。

## 【0063】

また、拡張部の先頭から一定時間経過した位置から再生開始することを指示するために、目標再生時間TargetTimeがmoof解析手段moofDemuxに入力される。なお、moov解析後にTargetTimeが決定する際には、メディア情報MoovInfに目標再生時間を含め、目標再生時間TargetTimeを別途入力することはしないものとしてもよい。例えば、拡張部の先頭から5s経過した位置から再生を開始したいときには、目標再生時間TargetTimeを5sに設定する。moof解析手段moofDemuxは、先頭のmoofから順に、moofに含まれるtrafのデータを分離し、trafデータTrafをtraf解析手段trafDemuxに入力する。traf解析手段trafDemuxは、各moofのtrafに含まれるサンプルの情報を解析し、解析結果trafInfをRA検索手段RAfindに入力する。解析結果trafInfは、trafに含まれるサンプルのサイズ、再生時間長、および、ビデオサンプルの場合には、画面内符号化フレームであるかどうか



かを判別するための情報を含む。

【0064】

RAfindは、拡張部において目標再生時間TargetTimeより後で、最も早い表示開始時刻をもつランダムアクセス可能なサンプルを検索し、検索されたサンプルがメモリMem1のどこに格納されているかを示すポインタ情報RAinfをサンプル取得手段GetSplに入力する。ランダムアクセス可能なビデオサンプルを検索する際には、moofにおける先頭trafの先頭trunに含まれる先頭サンプルが画面内符号化フレームであるかのみを検索するものとする。画面内符号化フレームであるかどうかは、先頭サンプルのsample\_flagsの値により判定する。なお、first\_sample\_flagsフィールドが存在する際には、first\_sample\_flagsの値のみを参照することにより判定してもよい。サンプル取得手段GetSplは、ポインタ情報RAinfに基づいてメモリMem1からサンプルデータsplDat1を取得し、splDat2として出力する。

【0065】

【発明の効果】

本発明の請求項1に係る多重化装置により出力されるデータによれば、パケットに含まれる各メディアの表示開始時刻、あるいはランダムアクセス可能なサンプルの配置を考慮してパケット化されているため、再生時にサンプルを検索する際の計算量を削減することができる。

【0066】

本発明の請求項2に係る多重化装置により出力されるデータによれば、パケットに含まれる各メディアの表示開始時刻が揃っているため、再生時にサンプルを検索する際の計算量を削減することができる。

【0067】

本発明の請求項3に係る多重化装置により出力されるデータによれば、パケットに含まれる各メディアの表示開始時刻が揃っているため、再生時にサンプルを検索する際の計算量を削減することができる。

【0068】

本発明の請求項4に係る多重化装置により出力されるデータによれば、隣接するパケットに含まれるメディアの表示開始時刻が揃っているため、再生時にサンプルを検索する際の計算量を削減することができる。

【0069】

本発明の請求項5に係る多重化装置により出力されるデータによれば、パケット内にビデオの画面内符号化フレームが含まれる際には、パケットの先頭サンプルが画面内符号化フレームとなるため、パケットの先頭サンプルから再生を開始でき、ビデオの画面内符号化フレームを検索する際の計算量を削減することができる。

【0070】

本発明の請求項6に係る逆多重化装置によれば、パケットに含まれる各メディアデータがランダムアクセス可能であるかどうかを調べる際に、各メディアの先頭サンプルのみ検索することにより判定可能となるため、ランダムアクセス時にサンプルを検索する際の計算量を削減することができる。

【0071】

本発明の請求項7に係る多重化方法を用いて出力されるデータによれば、パケットに含まれる各メディアの表示開始時刻、あるいはランダムアクセス可能なサンプルの配置を考慮してパケット化されているため、再生時にサンプルを検索する際の計算量を削減することができる。

【0072】

本発明の請求項8に係るデータ記憶媒体によれば、パケットに含まれる各メディアの表示開始時刻、あるいはランダムアクセス可能なサンプルの配置を考慮してパケット化する処理を、コンピュータに行わせるための処理を格納したので、該当プログラムをコンピュータにロードすることにより、ランダムアクセス時にサンプルを検索する際の計算量を削減可能なデータを出力する多重化装置を実現できる。

【図面の簡単な説明】

【図1】

本発明の実施の形態1および実施の形態2の多重化装置の構成を示すブロック

図

【図 2】

本発明の実施の形態 1 のパケット単位決定手段 `P a c k t i z e 2` の動作を示すフローチャート

【図 3】

本発明の実施の形態 1 の多重化装置が出力する `m p 4` データのデータ構造図

【図 4】

本発明の実施の形態 2 のパケット単位決定手段 `P a c k t i z e 2` の第 1 の動作を示すフローチャート

【図 5】

本発明の実施の形態 2 のパケット単位決定手段 `P a c k t i z e 2` の第 2 の動作を示すフローチャート

【図 6】

本発明の実施の形態 2 の多重化装置が出力する `m p 4` データのデータ構造図

【図 7】

本発明の実施の形態 3 の逆多重化装置の構成を示すブロック図

【図 8】

`MP 4` の `B o x` 構造を説明するための図

【図 9】

`MP 4` ファイルの基本部の構造を説明するための図

【図 10】

`m o o v` の構造を説明するための図

【図 11】

`MP 4` ファイルの拡張部の構造を説明するための図

【図 12】

`m o o f` と `t r a f` の構造を説明するための図

【図 13】

拡張部を使用したコンテンツの格納例を説明するための図

【図 14】

従来の多重化装置の構成を示すブロック図

【図15】

従来の多重化装置のパケット単位決定手段 `Pack t i z e 1` の動作を示すフローチャート

【図16】

従来の多重化装置のメモリ `Mem 2` に格納されるデータテーブルの構造を示す図

【図17】

従来の多重化装置が出力するMP4データのデータ構造を示す図

【図18】

`t r u n` の構造を示す図

【符号の説明】

`I n p u t 1`, `I n p u t 2` 入力手段

`Mem 1`, `Mem 2`, `Mem 3` メモリ

`P a r s e 1`, `P a r s e 2` 解析手段

`P a c k t i z e 1`, `P a c k t i z e 2` パケット単位決定手段

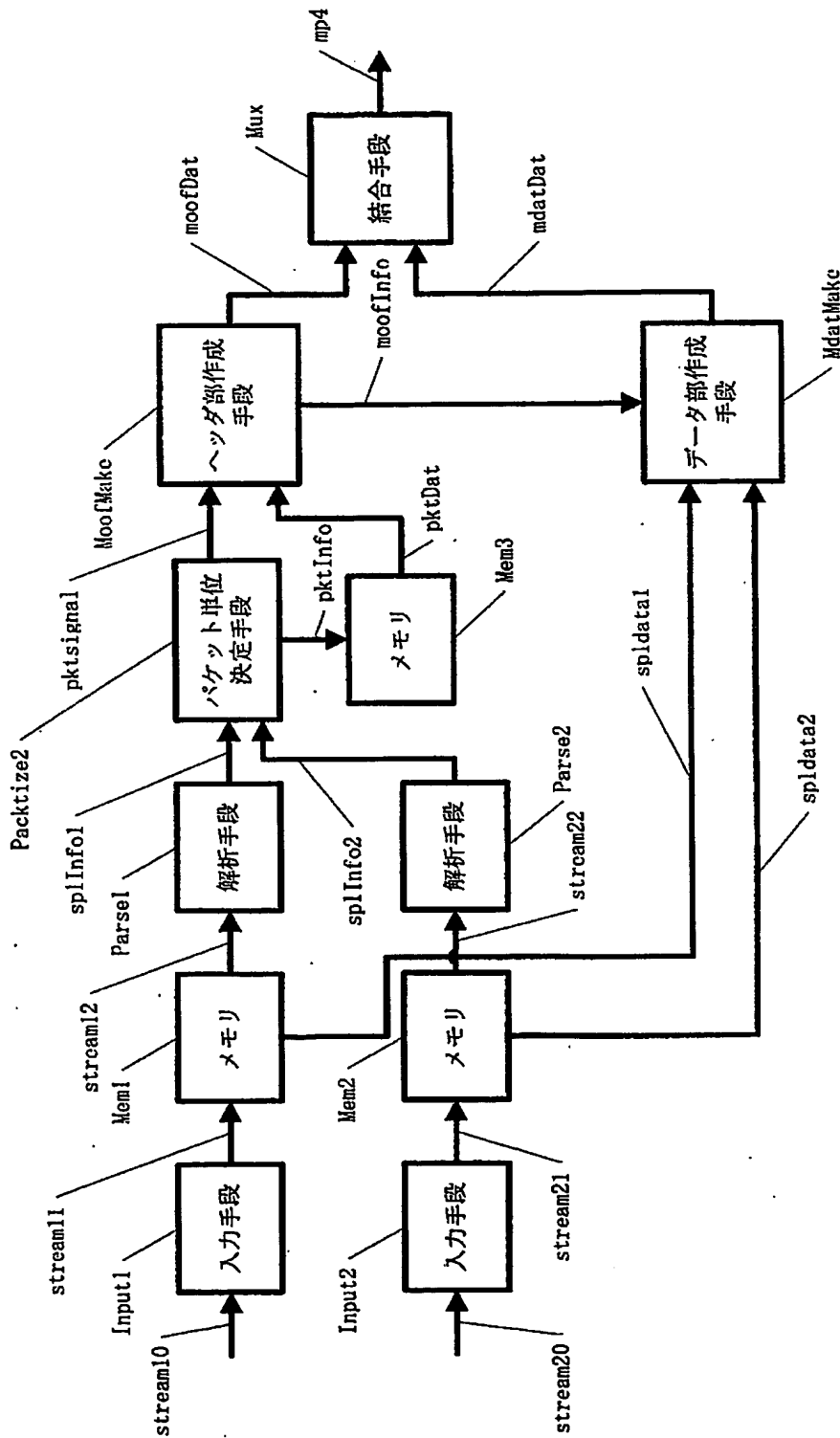
`M o o f M a k e` ヘッダ部作成手段

`M d a t M a k e` データ部作成手段

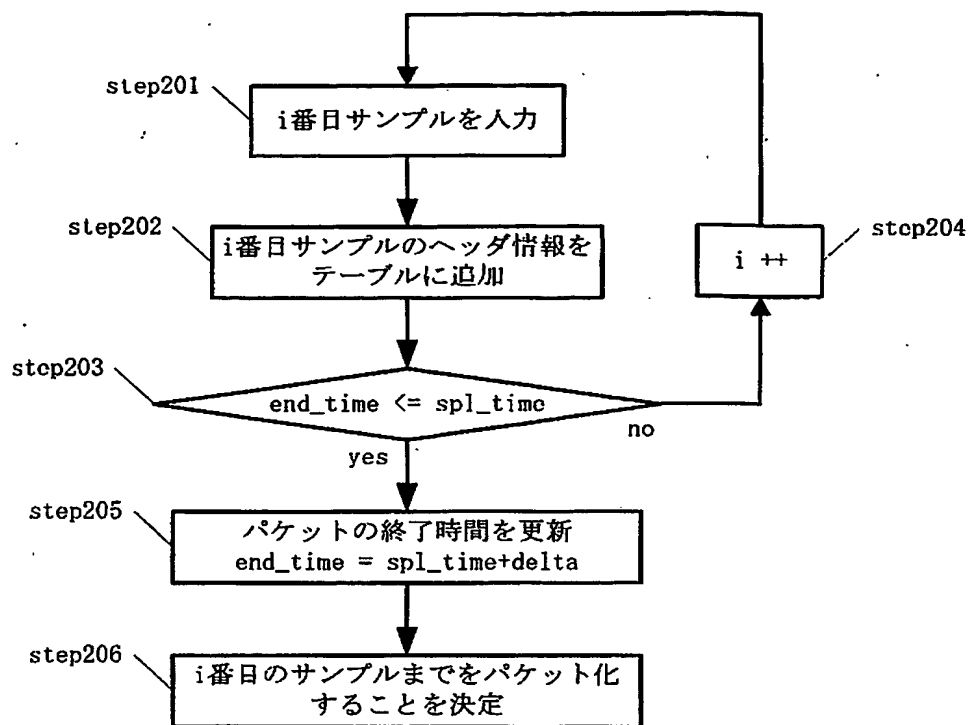
`M u x` 結合手段

【書類名】 図面

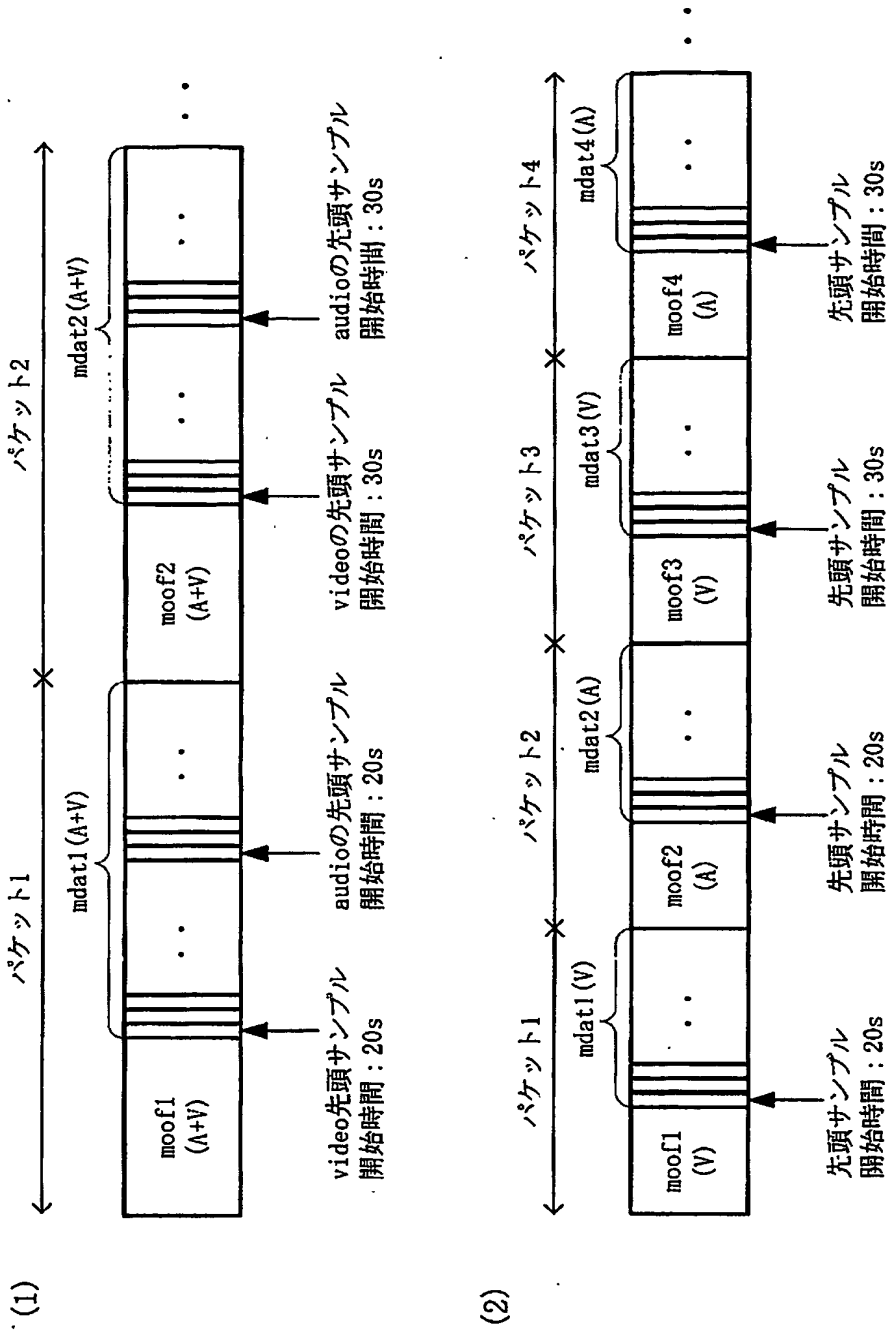
【図1】



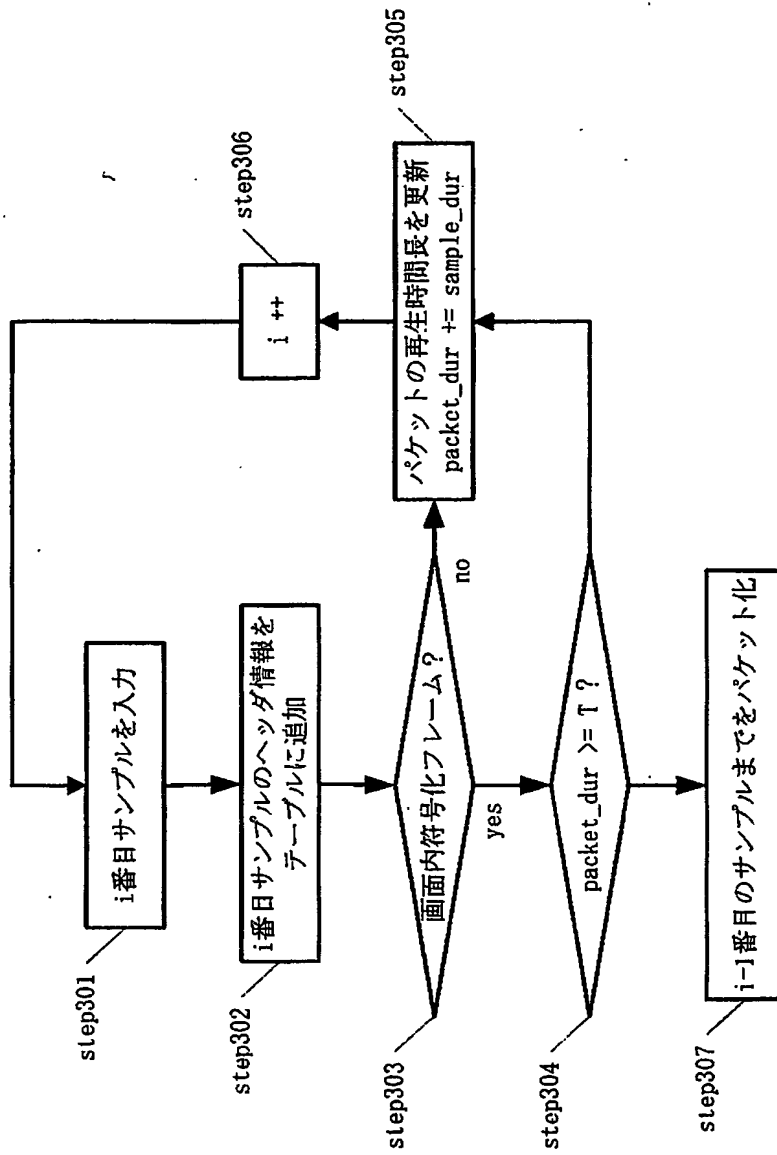
【図 2】



【図3】

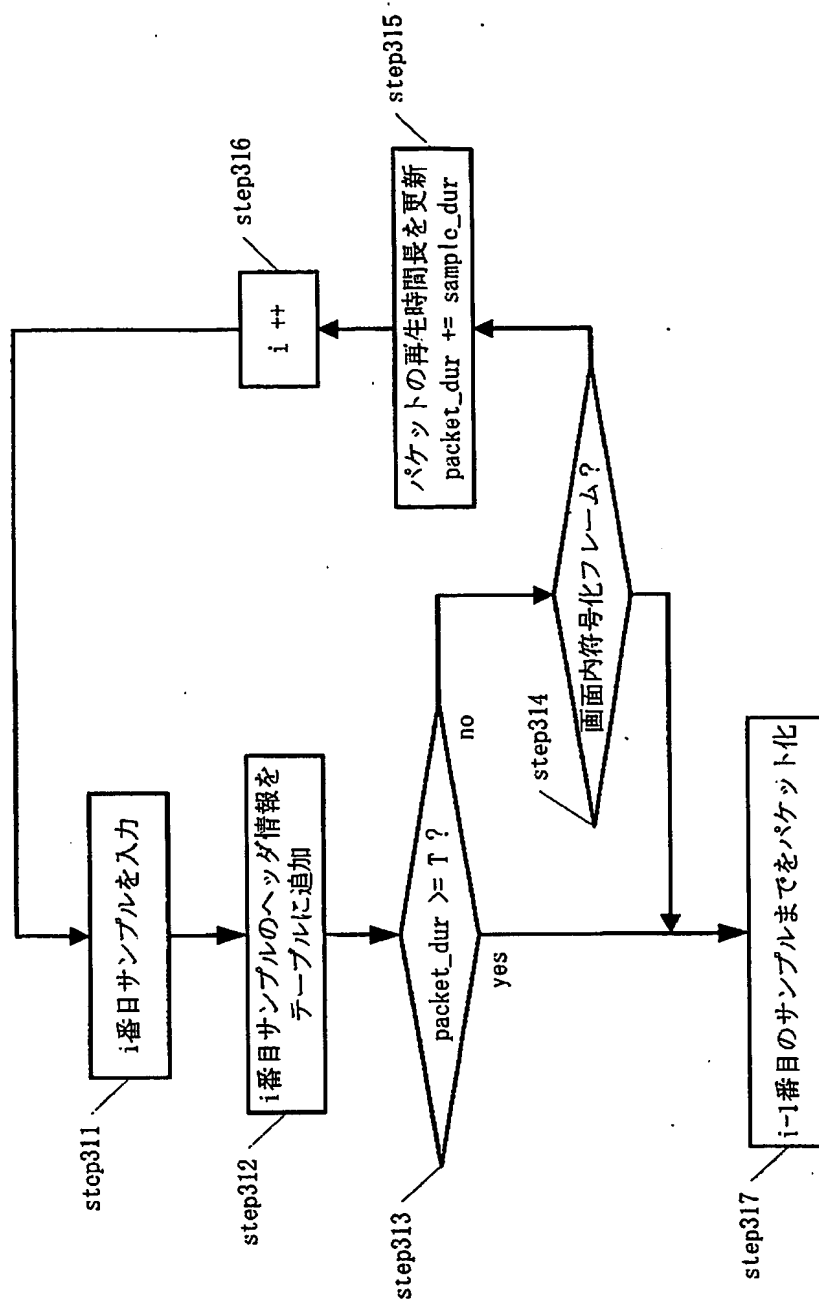


【図 4】

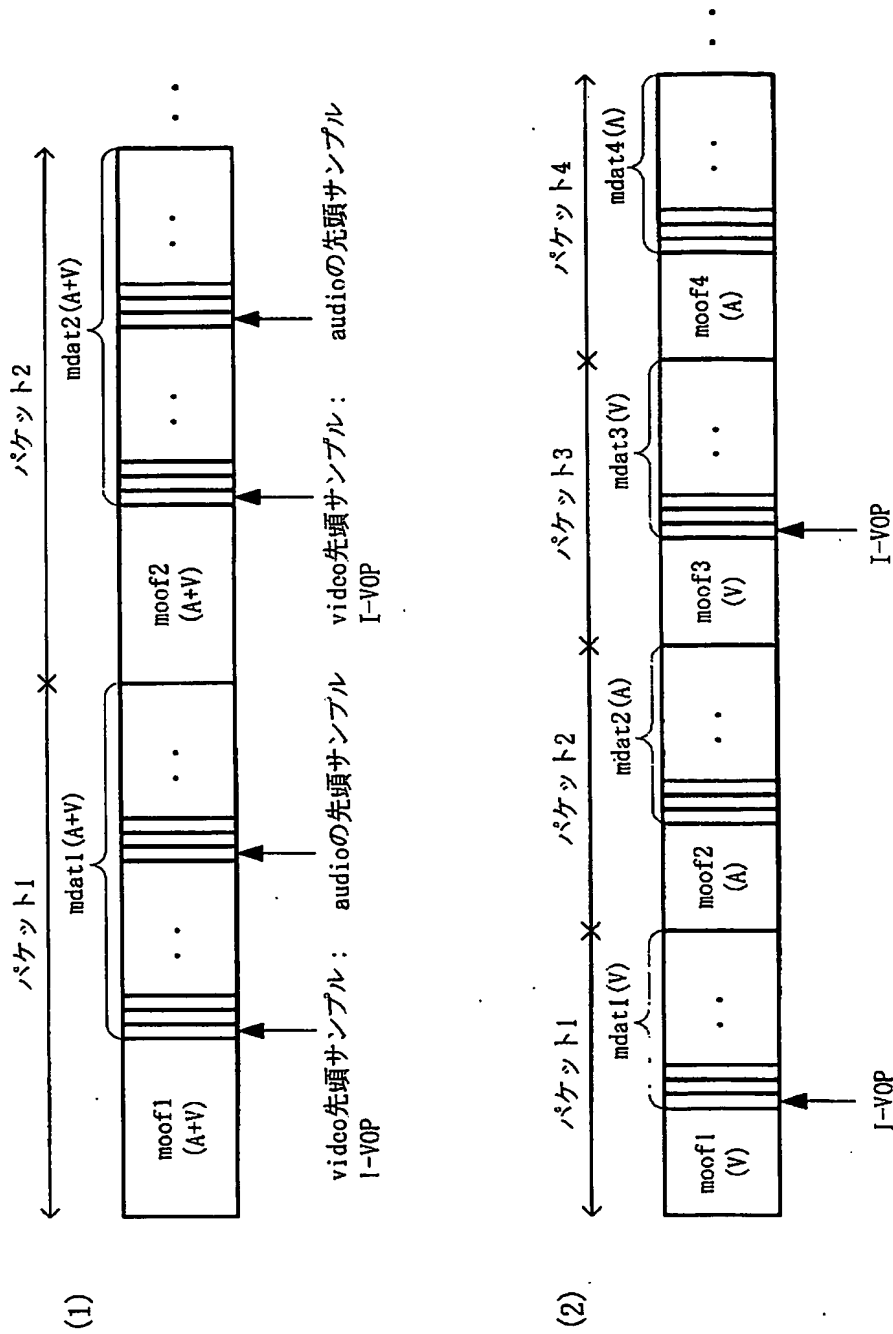




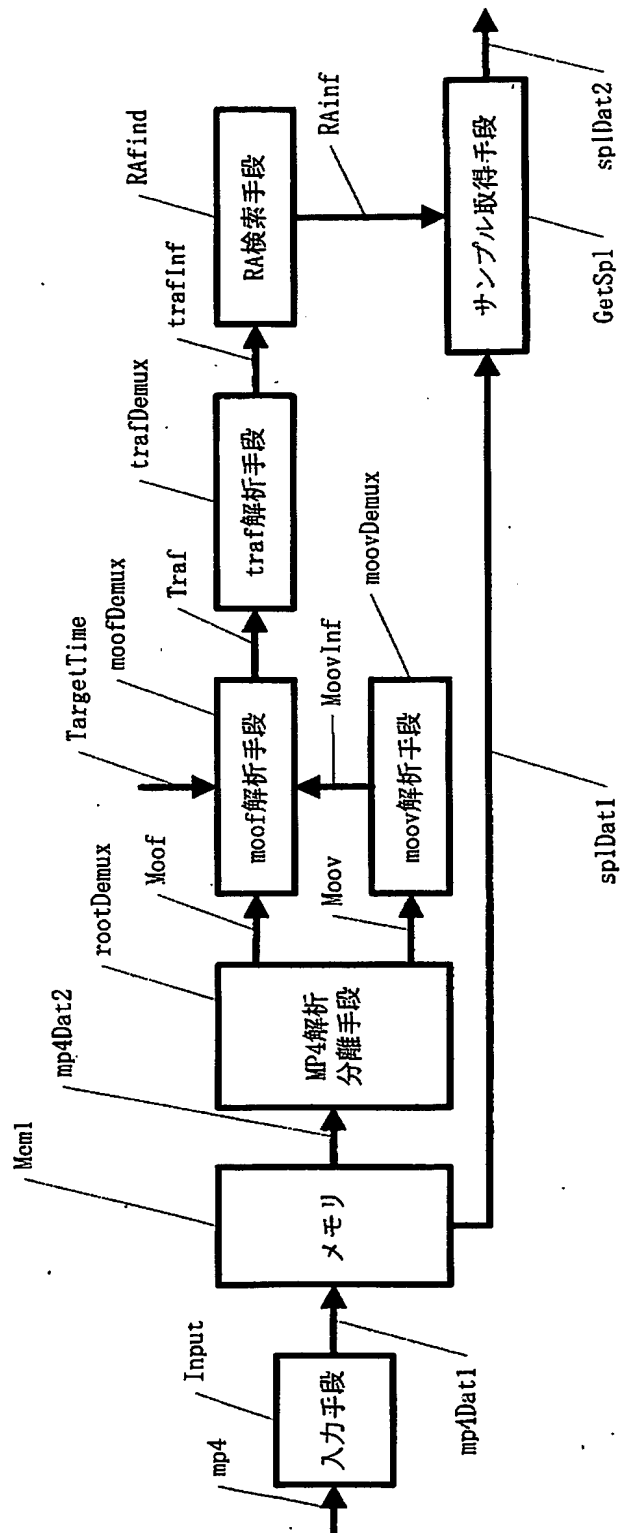
【図5】



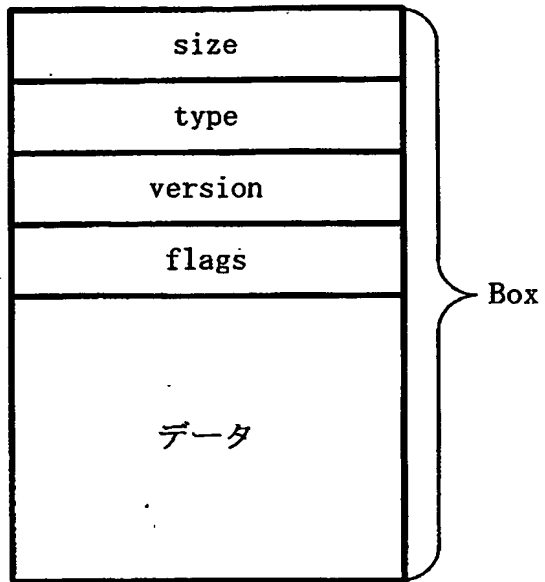
【図 6】



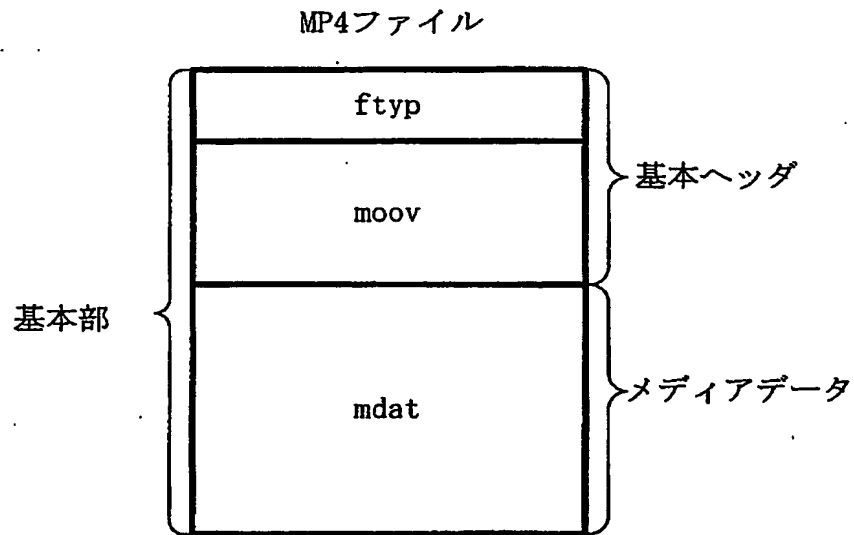
【図7】



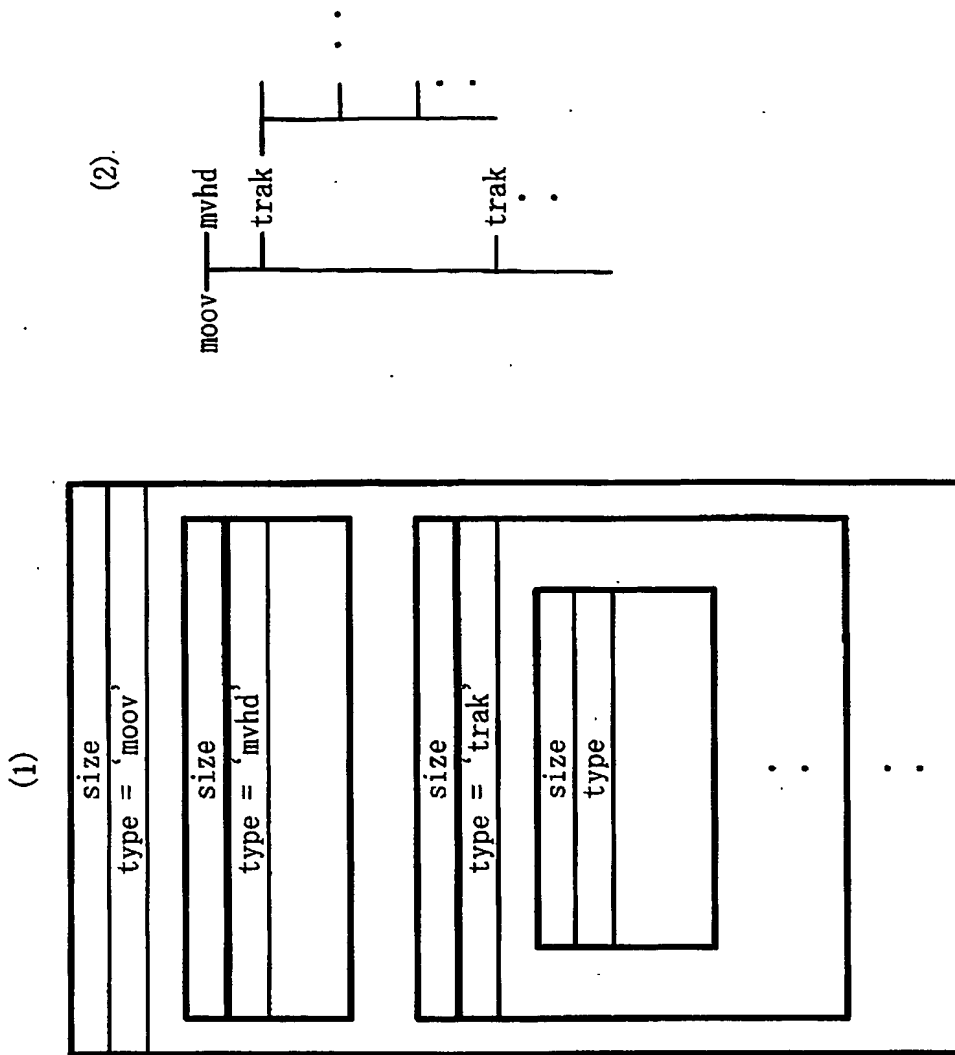
【図 8】



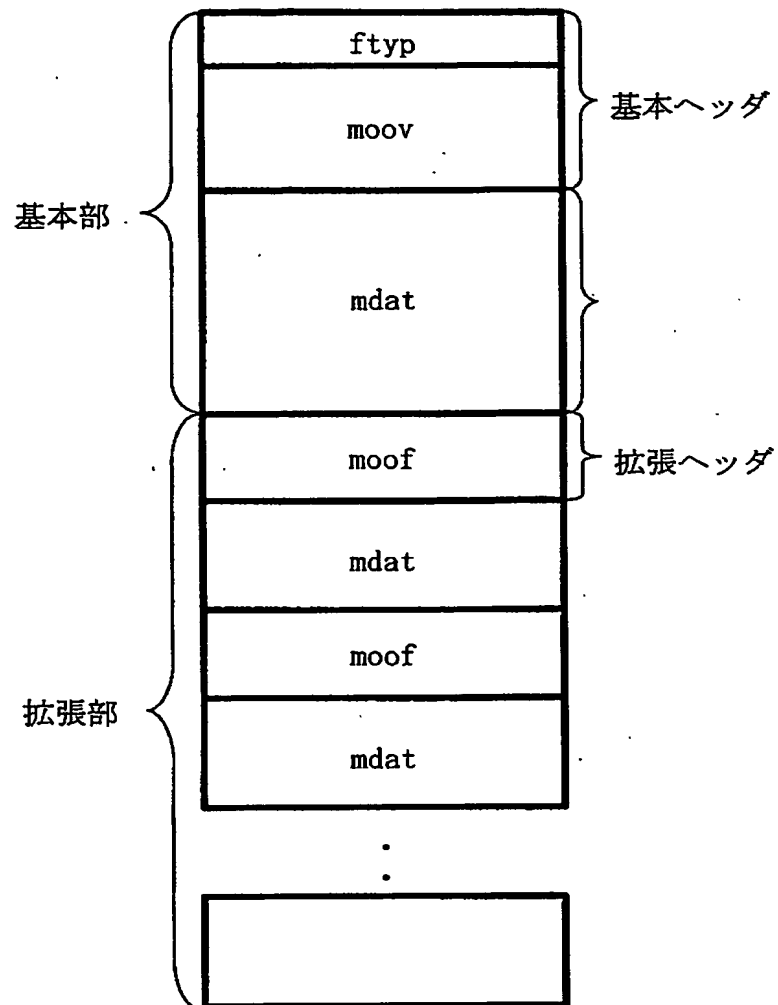
【図 9】



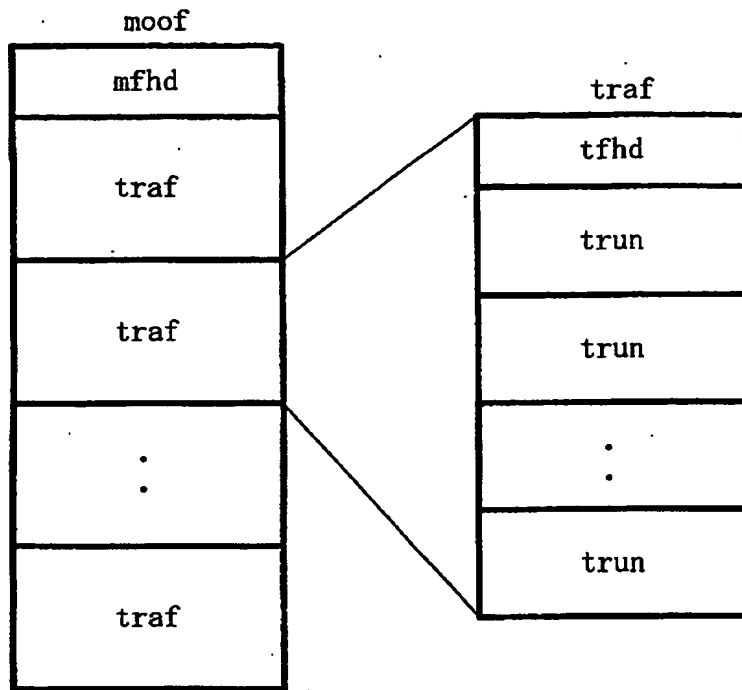
【図 10】



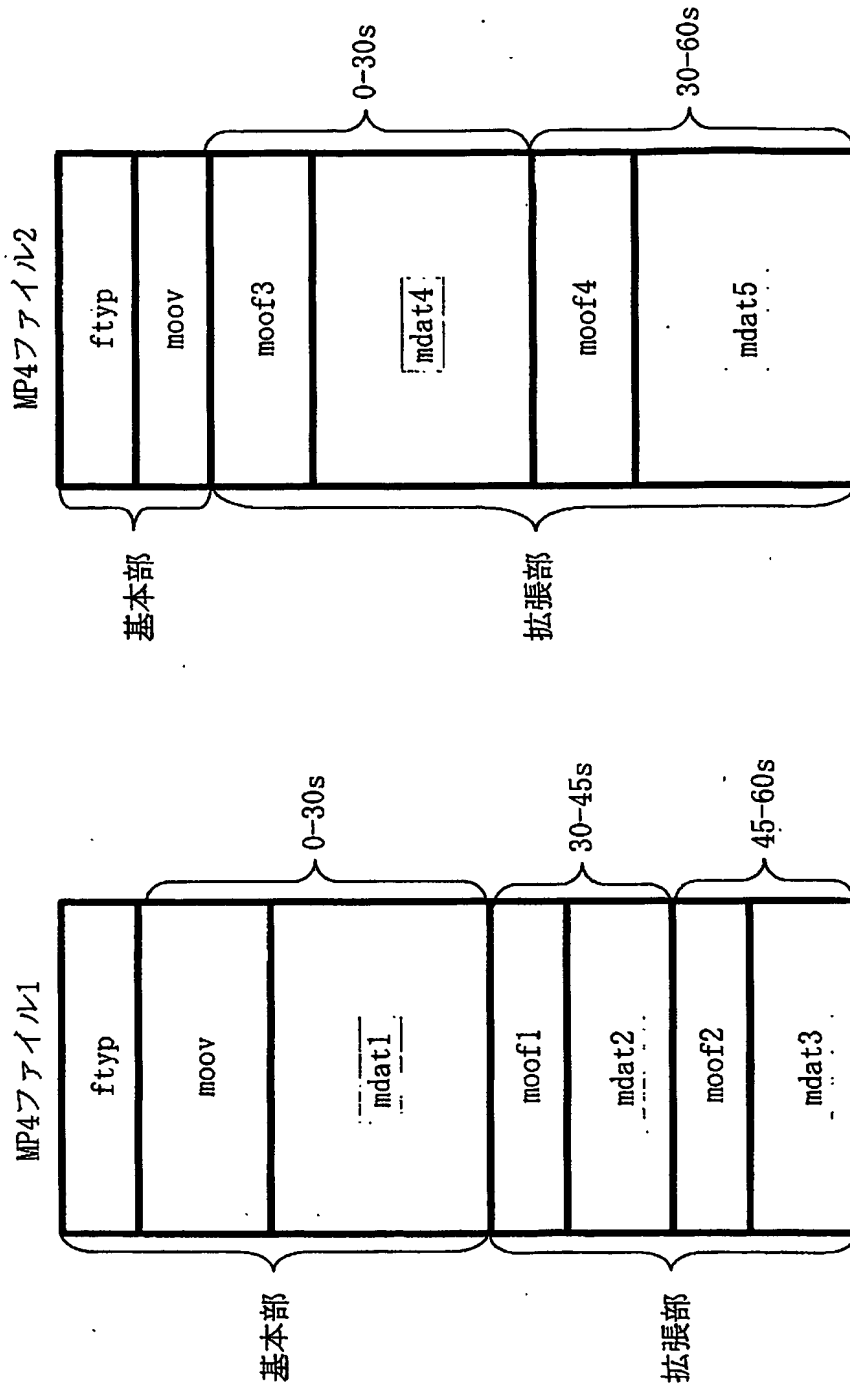
【図 11】



【図 12】

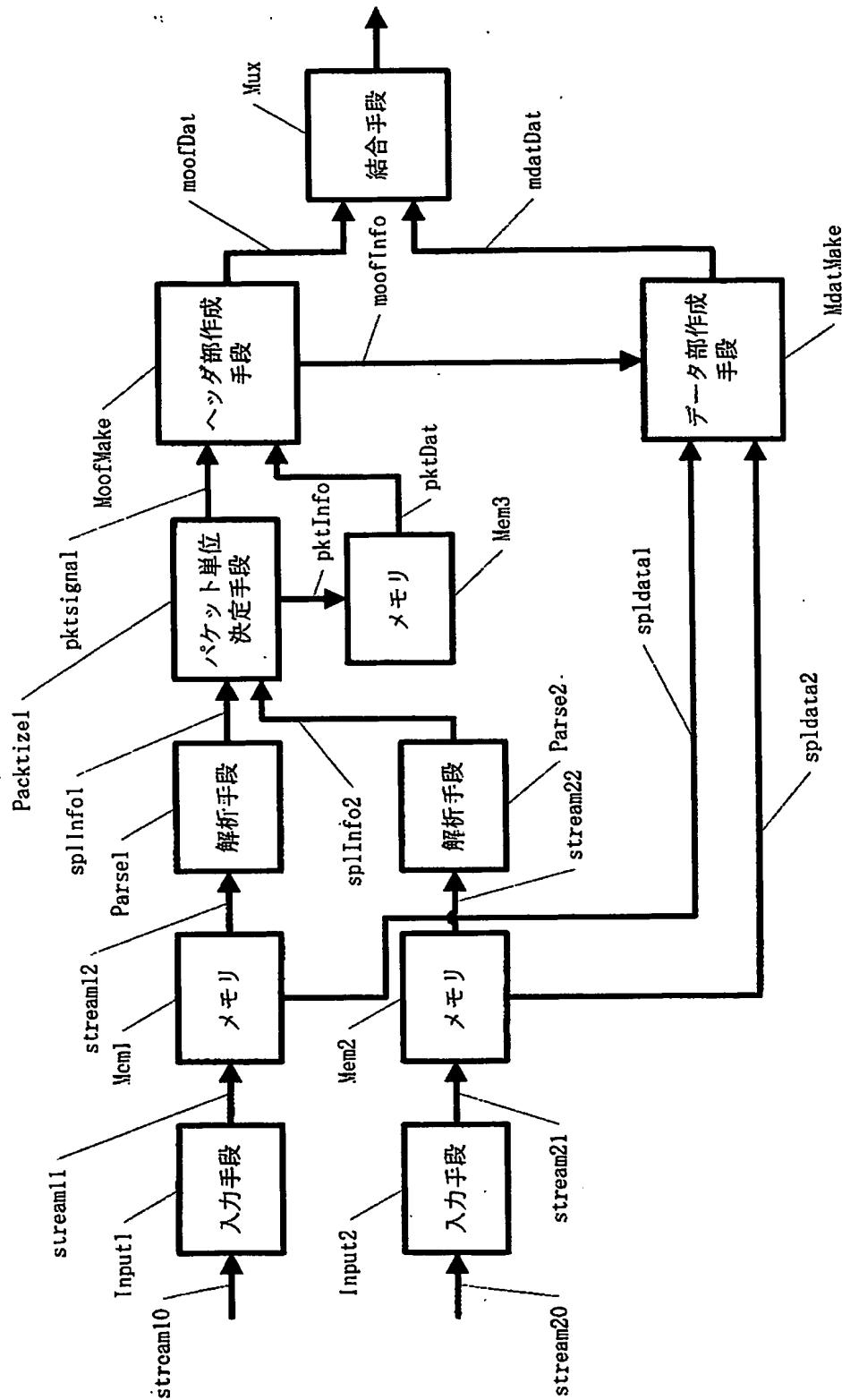


【図 13】

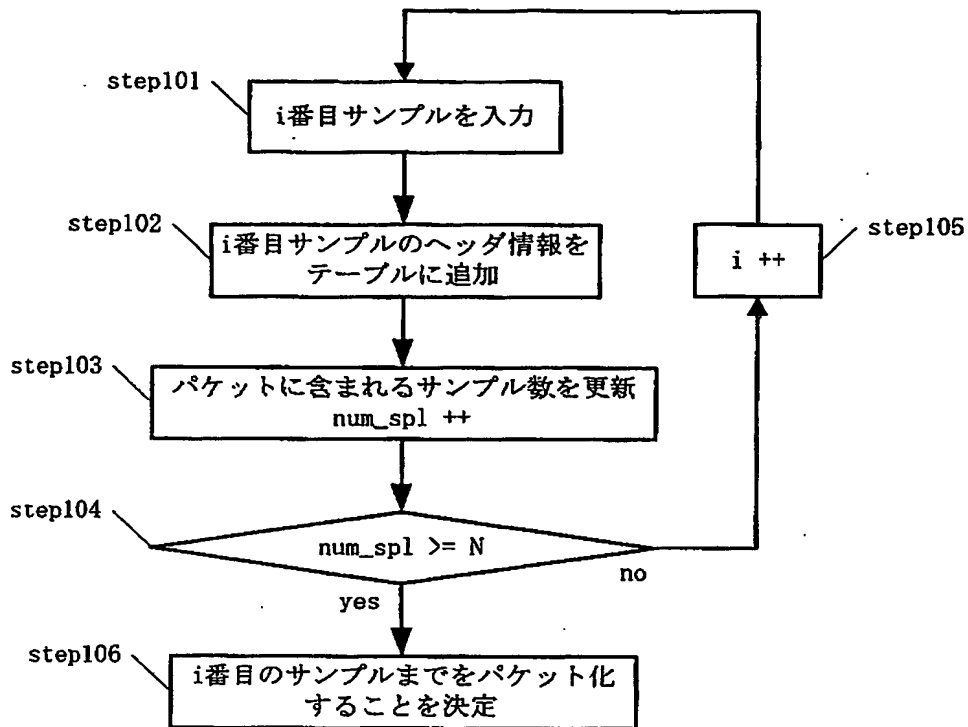




【図 14】



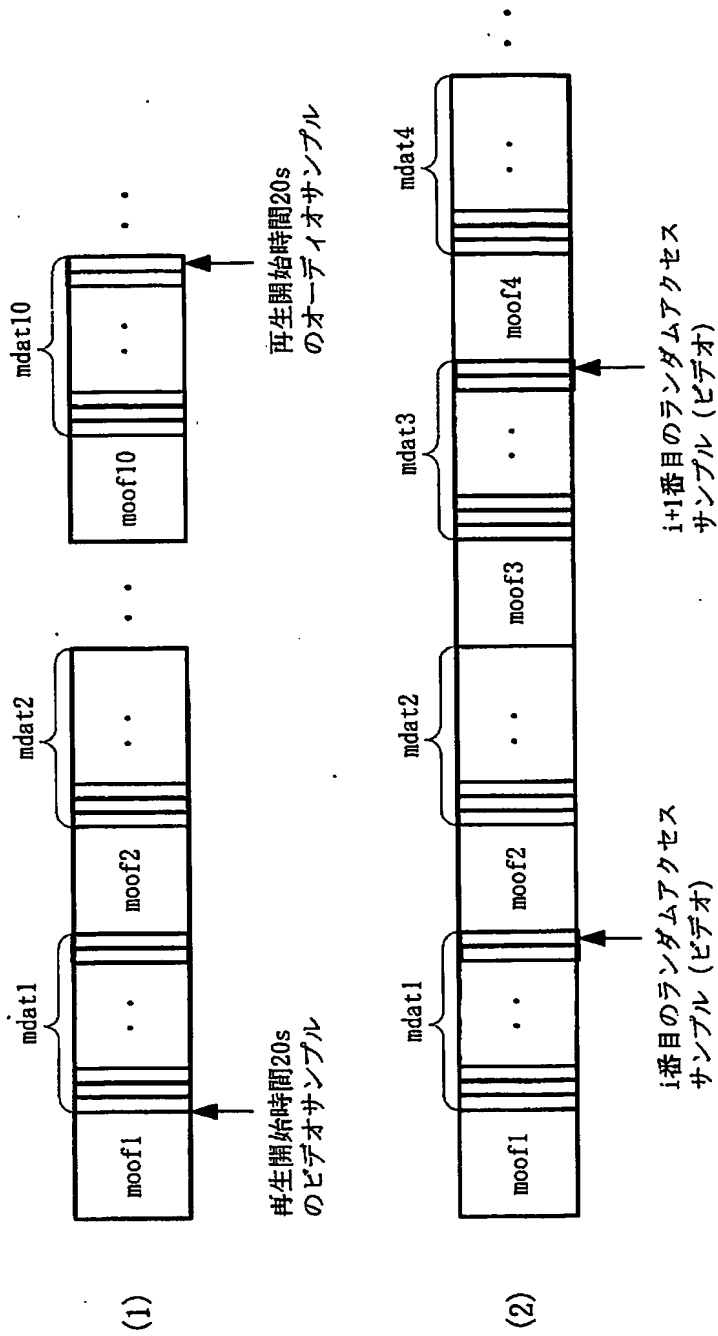
【図 15】



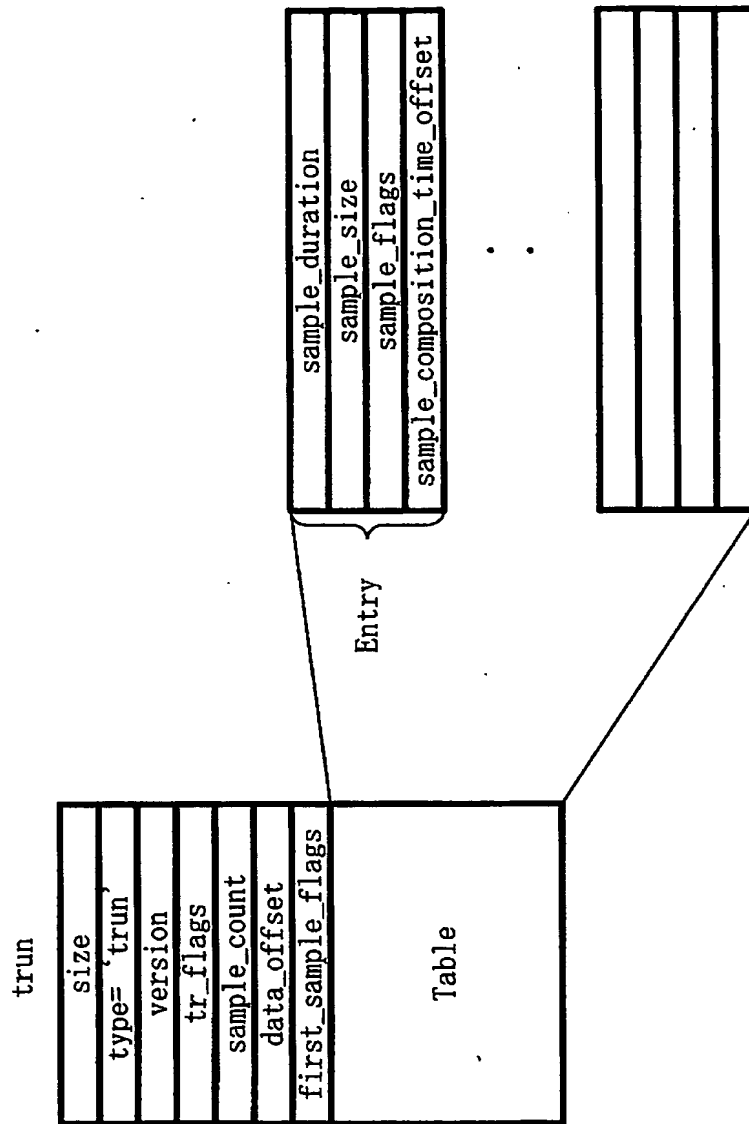
【図 16】

	サンプルのサイズ	サンプルの 再生時間長	サンプルは 画面内符号化フレーム？
サンプル1			
サンプル2	300	30	1
サンプル3			

【図17】



【図 18】



【書類名】 要約書

【要約】

【課題】 パケットのデータを途中から再生する際に、サンプルへのアクセスに必要な計算量が削減される拡張部のデータを出力する多重化装置を提供する。

【解決手段】 パケット単位決定手段 *P a c k t i z e 2* は、パケットに含まれる各メディアの先頭サンプルの再生開始時間が揃うように、あるいは、パケット内の動画データが画面内符号化フレームのデータを含む際には、パケットに含まれる動画データの先頭データが画面内符号化フレームのデータとなるようにパケット単位を決定する。

【選択図】 図 1

特願2002-185758

出願人履歴情報

識別番号

[000005821]

1. 変更年月日

1990年 8月28日

[変更理由]

新規登録

住 所

大阪府門真市大字門真1006番地

氏 名

松下電器産業株式会社